Pricing a Bermudan Option under the Constant Elasticity of Variance model

Kwaku Quinton Rwexana

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

October 26, 2017

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy in the University of the Cape Town. It has not been submitted before for any degree or examination in any other University.

October 26, 2017

Abstract

This dissertation investigates the computational efficiency and accuracy of three methodologies in the pricing of a Bermudan option, under the constant elasticity of variance (CEV) model. The pricing methods considered are the finite difference method, least squares Monte Carlo method and recursive marginal quantization (RMQ) method. Specific emphasis will be on RMQ, as it is the most recent method. A plain vanilla European option is initially priced using the above mentioned methods, and the results obtained are compared to the Black-Scholes option pricing formula to determine their viability as pricing methods. Once the methods have been validated for the European option, a Bermudan option is then priced for these methods. Instead of using the Black-Scholes option pricing formula for comparison of the prices obtained, a high-resolution finite difference scheme is used as a proxy in the absence of an analytical solution. One of the main advantages of the recursive marginal quantization (RMQ) method is that the continuation value of the option is computed at almost no additional computational cost, this with other contributing factors leads to a computationally efficient and accurate method for pricing.

Acknowledgements

I thank my supervisor Professor Thomas McWalter for his incredible insight and motivation as well as my co-supervisor Ralph Rudd for showing me how to think when a task seemed insurmountable (also for lending me your laptop). I thank Professor David Taylor for his encouragement, support and kind words throughout my journey. I thank Lameez Fakier for solving all my administration related issues (even on short notice). I thank my family Nokuthula Rwexana, Noxolo Mshumpela, Malusi Mshumpela and Nokuthula Tana for your unending support and understanding throughout my life in the pursuit of my goals. Finally I would like to thank Liberty Financial Solutions for making this journey possible, in providing me with financial aid for the 2016 academic year.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The derivatives market is like any other financial market, in that the participants in the market and the reasons they use it are varied. The participants of the derivatives markets may be broadly divided into three categories. The first are those that participate in the market to mitigate present or anticipated exposure to the underlying asset because of the uncertainty of the price. These participants (organization, corporates, individuals, etc.) make an investment in derivative instruments to offset the potential loss of their investment, this is also termed as hedging their risk. The second take a view on the future direction of prices, whether prices will go up or down, and hence either buy or sell derivative instruments. These participants are known as speculators. Then, the third take positions in derivative instruments to exploit market inefficiencies in order to earn a riskless profit (a so called arbitrage), if they are successful, then this activity sets the price back into equilibrium.

Derivative instruments serve many purposes but the three most common are the mitigation of risk, speculative activity and the discovery of price. Some analysts estimate that the derivatives market as reported by the Bank for International Settlement (BIS) as of June 2017, amounted to approximately 33 trillion USD. Although derivative instruments play a crucial role in the economy, there is also the risk that these instruments may exacerbate market stresses.

The standard definition of an option contract is the right, without the obligation, to buy or sell the underlying asset (share, gold, oil, etc.) at a later date under certain conditions. In the case of a call option, it is the right to buy the underlying asset and in the case of a put option, it is the right to sell the underlying asset without the obligation to do so. The use of these contracts is not a modern development but arose from the process of exchange in markets. The option price is determined by a combination of two factors, time value (extrinsic value) and intrinsic value.

option value = intrinsic value + time value.

The intrinsic value for an option is the difference between underlying asset and the strike price. For a call option,

underlying asset (S) - strike price (K),

and for a put option,

strike price (K) - underlying asset (S).

The strike price is the fixed price where the holder of the call option has the right to buy and the holder of a put option has the right to sell. The time value (extrinsic value) is the benefit received for the possibility of the underlying asset moving into the money where, $S > K$ for a call option and $K > S$ for a put option. Black and Scholes (1973) derived the basic model for calculating the theoretical price of an option. The model assumes that the option can only be exercised at expiration, dividends are not paid out during tenure of the option, no commissions are paid, the volatility and risk free rate are constant, the returns of the underlying are normally distributed and markets are efficient. In order to value the option according to this model, the current underlying price, the strike price, the time to expiration, the implied volatility and the risk free rate are required.

The Black and Scholes (1973) pricing formula for a call option is then defined to be,

$$C = SN(d_1) - N(d_2)Ke^{rt},$$

where

$$d_1 = \frac{ln\frac{S}{K} + \frac{r+\sigma^2}{2}t}{\sigma\sqrt{t}},$$

and

$$d_2 = d_1 - \sigma\sqrt{t}.$$

The first part, $SN(d_1)$, is the price of the underlying asset multiplied by the change in the call premium, while the second part, $N(d_2)Ke^{rt}$, is the current value of paying the exercise price at expiration.

When considering options, two extremes are possible. The first being a European option which only allows exercise of the option at the terminal time and the second being the American option which is characterized by the possibility of early exercise at any time. Bermudan options have behaviour that is intermediate between

American and European options in terms of their exercise. Unlike the European option, the Bermudan option can be exercised at more than one date, but in contrast to the American option, the exercise dates are finite in number. Therefore, Bermudan options are "in-between" American and European options, hence the reference to Bermuda.

The difficulty with an option that has an early exercise feature is that it is not known beforehand when exercise of the option will maximize the payoff. This is known as the optimal stopping problem as described by El Karoui and Karatzas (1995). McKean (1965) first proposed that this can be overcome by determining the option price and early exercise boundary simultaneously. There is, however, no known analytical solution for such a problem, even when the option has a simple payoff and underlying asset dynamics. This problem must be solved by implementing numerical methods.

1.1 Numerical methods

Many numerical methods have been introduced in the past, this includes the finite difference scheme pioneered by Schwartz (1977) and Brennan and Schwartz (1978) and the binomial model introduced by Cox *et al.* (1979). These methods are convergent and simple but are not computationally efficient.

Barone-Adesi and Whaley (1987) proposed the first analytical approximations for American options, using what is known as the quadratic method of MacMillan (1986). The method is highly efficient and accurate but has difficulty converging under certain conditions. Others, like Johnson (1983) and Broadie and Detemple (1996), have provided bounds (upper and lower) for options with an early exercise feature, which are based on regression coefficients.

Sullivan (2000), approximated the option value function using Chebychev polynomials and also incorporated the Guassian quadrature integration scheme at each discrete exercise date. Unfortunately the convergence properties of this numerical scheme are unknown. Geske and Johnson (1984) approximated American options using a truncated series of multivariate normal distributions, the convergence in this scheme is obtained by adding more terms to the series as the accuracy of the scheme is a function of the number of terms. A Richardson extrapolation is also employed to increase efficiency but the scheme may suffer from numerical instability.

Carr (1998) uses a randomization approach which is faster and offers better accuracy but also uses the Richardson extrapolation scheme. Kim (1990), Jacka (1991), Carr *et al.* (1992) and Jamshidian (1989) propose a new method which employs integral representations to describe the early exercise premium.

In this dissertation, three numerical methods are implemented. The first being the finite difference method, pioneered by Schwartz (1977) and Brennan and Schwartz (1978), this method is accurate but not computationally efficient, especially when the resolution of the grid size is increased to ensure accuracy. The second method is the least squares Monte Carlo method of Longstaff and Schwartz (2001), this approach combines Monte Carlo simulation with a least squares regression and, in general, this method provides a good estimate of the option value but does not compare well in terms of accuracy to the Finite Difference method (Tompaidis and Yang, 2014). The third method is recursive marginal quantization (RMQ), which is a new approach proposed by Pagès and Sagna (2015) and extended by McWalter *et al.* (2017), this method is proposed to be computationally efficient and accurate.

1.2 Stochastic processes and partial differential equations

The stochastic process for asset price movements that are considered normal, meaning very few market shocks occur is known as geometric Brownian motion. This process is assumed to follow a Markov process in that only the present value of the asset price is relevant for the future price movements. This assumes that the present value of the asset price contains all price history and anything before the present moment is irrelevant. The price movements of the asset can be described by the following Itô process or stochastic differential equation,

$$dS_t = m(t, S_t)dt + \sigma(t, S_t)dW_t,$$

where S_t is the current underlying price, $m(t, S_t)$ is the drift, $\sigma(t, S_t)$ is the annualized instantaneous volatility of the underlying asset and W_t the Wiener process. Black and Scholes (1973) assume that the above equation takes the form of $m(t, S_t) = \mu S_t$, $\sigma(t, S_t) = \sigma S_t$ and thus the asset price return follows,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

known as geometric Brownian motion. The first term μdt is the average return over a short time interval dt and the second term σdW_t is known as the noise or the random changes of the asset price to external factors such as news.

Bru and Yor (2002) state that, according to Itô's Lemma, if f is a function of two variables S, t then this function can be expressed as,

$$df = \sigma S \frac{df}{dS} dW + (\mu S \frac{df}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 f}{dS^2} + \frac{df}{dt}).$$

Let $V(S,t)$ be the value of an option, let r be the interest rate, σ the instantaneous volatility and μ the drift. Using Itô's lemma, the value of the option is expressed as,

$$dV = \sigma S \frac{dV}{dS} dW + (\mu S \frac{dV}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 V}{dS^2} + \frac{dV}{dt})t.$$

Now, consider a portfolio containing one option and -Δ units of the underlying asset. The value of the portfolio is,

$$\Pi = V - \Delta S.$$

Then the incremental change in value of the portfolio is

$$d\Pi = dV - \Delta dS.$$

Then substitute the value of the option

$$d\Pi = \sigma S \frac{dV}{dS} dW + (\mu S \frac{dV}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2 V}{dS^2} + \frac{dV}{dt})dt - \Delta S \mu dt - \Delta S \sigma dt.$$

Choose $\Delta = \frac{dV}{dS}$ to get,

$$d\Pi = (\frac{1}{2}\sigma^2 S^2 \frac{d^2 V}{dS^2} + \frac{dV}{dt})dt.$$

Now if Π was invested in a riskless asset it would see a growth of $r\Pi dt$ in the interval of length dt. Then for a fair price we should have

$$r\Pi dt = (\frac{1}{2}\sigma^2 S^2 \frac{d^2 V}{dS^2} + \frac{dV}{dt})dt.$$

This can now be simplified to

$$r(V - \frac{dV}{dS}S) = \frac{1}{2}\sigma^2 S^2 \frac{d^2 V}{dS^2} + \frac{dV}{dt}.$$

In this process the Black and Scholes (1973) partial differential equation is then derived

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

This finding is important as it connects the stochastic process, in this case geometric Brownian motion to the partial differential equation of Black and Scholes (1973), which will be used extensively in this dissertation.

1.3 Research method and aims

This dissertation uses the work mentioned above and extends it, by pricing a Bermudan option under the constant elasticity of variance (CEV) model. The main focus of the dissertation will be on using recursive marginal quantization (RMQ) proposed by Pagès and Sagna (2015) and extended by McWalter *et al.* (2017) for higher order schemes to price a Bermudan option and compare this to the finite difference method and least squares Monte Carlo method to analyse the computational efficiency and accuracy of the three methods.

Since Bermudan options do not have a closed-form solution, analysis is initially done on pricing a European option, as this type of option has a known analytical solution for both the GBM and CEV diffusion processes. In the case of GBM diffusion process, the European option prices obtained using the three methods will be compared against the Black and Scholes (1973) option pricing formula. In the case of CEV diffusion process, the prices will be compared to the analytical solution of Schroder (1989), which was reformulated for the non-central chi-squared distribution. Once the European options are priced and compared to the analytical solution for validation, the Bermudan option will be priced using the three methods, and accuracy will be determined by taking the absolute difference between the option prices obtained and a high-resolution Crank-Nicolson finite difference scheme, which will act as the proxy for the exact solution of the Bermudan option price.

The price of a Bermudan call option $C(S,t)$ is required to satisfy:

$$C(0,t) = 0 \qquad \text{for all } t \in [0,T) \qquad \text{and} \qquad C(S,T) = \max(S-K,0),$$

and the price of a Bermudan put option $P(S,t)$ is required to satisfy:

$$P(\infty, t) = 0 \qquad \text{for all } t \in [0, T) \qquad \text{and} \qquad P(S, T) = \max(K - S, 0).$$

Since early exercise is restricted to certain dates $t_1 < t_2 < t_3 < \cdots < t_{n+1}$ during the option life, where $t_1 = 0$ and $t_{n+1} = T$, where T is the maturity of the option. The following discrete boundary conditions are also required

$$C(S, t_i) = \max(C(S, t_i^+), S - K),$$

for a call option or,

$$P(S, t_i) = \max(P(S, t_i^+), K - S),$$

for a put option, for $i = 1, 2, \ldots, n + 1$.

Given the above, a call option on a stock that pays no dividends is never exercised early as shown by Hull (2006). Therefore, only the pricing of a European put option and a Bermudan put option will be considered.

This dissertation consists of four chapters. Following this, a brief review of the constant elasticity of variance model is given, giving justification as to why this model was chosen. Followed by the different pricing models that will be investigated with a specific emphasis on recursive marginal quantization (RMQ). Lastly the results are presented, followed by the conclusion. It was found that of the three methods investigated, the least squares Monte Carlo method was least accurate in pricing a Bermudan option while the RMQ weak order 2.0 Taylor scheme was found to be the most accurate and the most efficient method. The RMQ Euler-Maruyama scheme was found to be the method with the shortest execution time when cardinality was varied.

Chapter 2

The constant elasticity of variance (CEV) model

To help introduce this topic, consider a arbitrary example, call it cellular behavior. Let $c(t) = c_t$ be a deterministic differential equation denoting the cells present at any time t. Now for a small incremental change in time, dt, the change in the cells is defined by the equation below,

$$\mathbf{d}c_t = Y c_t dt$$
$$\mathbf{c}(0) = c_0,$$

where Y is a constant. If some degree of variability is introduced, where the initial condition c_0 can no longer be considered deterministic, then c_0 becomes a random variable $C_t(z)$. The differential equation modelling the cells is then

$$\mathbf{d}C_t(z) = Y C_t(z) dt$$
$$\mathbf{C}_0(z).$$

The solution to the equation above is then given by $C_t(z) = C_0(z)exp(Yt)$ which can be considered to be a diffusion process. The path these cells take is determined by what $C_0(z)$ is. To add a bit more complexity, if Y is not known for certain, but it is known that it changes by a degree of randomness, then this process can be modelled as a stochastic process by introducing dX_t for which $\{x_t(z), t \geq 0\}$, then the differential equation is given by

$$\mathbf{d}C_t(z) = (Y dt + dX_t(z))C_t(z)$$
$$\mathbf{C}_0(z),$$

where dX_t is the variability added to Y.

The general diffusion process above was shown to be the solution to a differential equation describing the presence of cells. Now consider two models for describing the underlying dynamics of an asset. The first being geometric Brownian motion (GBM). This is the simplest model commonly used for continuous-time asset pricing. The stochastic differential equation is

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \tag{2.1}$$

where the solution to the above equation is given by

$$S_t = S_0 exp((\mu - \frac{\sigma^2}{2})t + \sigma W_t),$$

where μ is the drift, σ is the variance, S_t is the underlying asset and W_t is a standard Brownian Motion.

According to this process, the relative change in the underlying asset is a combination of deterministic growth and a normally distributed random change. The main assumption of GBM is that the underlying asset price follows a lognormal diffusion process (Schroder, 1989), meaning, the log returns, over the interval Δt are normally distributed with mean, $(\mu - \frac{1}{2}\sigma^2)\Delta t$ and variance, $\sigma^2 \Delta t$.

One major criticism of GBM is that log-normality does not hold exactly, and that the volatility of log returns are not constant (Delbaen and Shirakawa, 2002). This has been shown in many empirical studies.

An improved model was suggested by Cox (1975) and Cox and Ross (1976) who proposed the constant elasticity of variance (CEV) model as a better process for modelling stock price paths than GBM, moreover it is analytically tractable. The SDE for the CEV model is

$$dS_t = \mu S_t \, dt + \sigma S_t^{\alpha} \, dW_t, \tag{2.2}$$

where α is the elasticity parameter of the volatility, μ is the drift and σ is a constant. The instantaneous (local) volatility is given in terms of σ by $\sigma_{LN}(S_t) = \sigma S_t^{\alpha-1}$. This shows that instantaneous volatility is related to the stock price, which is known as the leverage effect in the equity markets (Cox and Ross, 1976).

When $\alpha = 0$, the constant elasticity of variance model becomes the constant volatility geometric Brownian motion process. Cox (1975) initially investigated the case for which the volatility is a decreasing function of the asset price, meaning, as the stock price declines, the stock price volatility increases. This relationship between the volatility and the asset price gives the implied volatility skew. The elasticity parameter, α, controls the steepness of the skew, while σ, which is the scaling parameter, fixes the at-the-money volatility level.

The model has found empirical support in the literature, an example being the Bekaert and Wu (2000) study which found the existence of a negative correlation between stock returns and stock volatility. The inverse relationship between implied volatility and strike price of an option contract is observed by Dennis and Mayhew (2002), in the investigation of the prices of stock options traded on the Chicago Board Options Exchange.

The method of pricing European options under the CEV model has become well established since the work done by Cox (1975), but the same cannot be said for Bermudan options. The CEV model will be the diffusion process that is adopted for all underlying asset dynamics in this dissertation, unless stated otherwise.

Chapter 3

A review of the numerical methods

3.1 Finite difference method

From the previous chapter, it is known that a Bermudan put option does not have a closed-form solution, much like American put options, making pricing such an instrument a challenge. So instead, a European put option is priced initially to validate the methods. In this section the focus will be on the formulation of the finite difference method for pricing a European put option and how this method is then adapted for the Bermudan put option.

To introduce the idea, consider an example. The definition of a derivative is given by

$$\frac{df}{dx} = lim_{\Delta t \to 0} \frac{(f(x + \Delta t) - f(x))}{\Delta t}.$$

The value of $\frac{df}{dx}$ can also be approximated by using finite difference approximation,

$$\frac{f(x + \Delta t) - f(x)}{\Delta t}$$

over a small interval, Δt.

The idea presented above is to show that differential equations can be solved numerically, by replacing the derivatives in the equation with a finite difference approximation on a discretised space. The main assumptions are that f is a smooth function and is continuously differentiable sufficiently many times and that the domain the function is described over can be discretised with a uniform grid.

The basic idea of the finite difference approach used in this section is to solve a partial differential equation (PDE), by replacing the partial derivatives with approximations obtained by a Taylor series expansion about a certain value (Strikwerda, 2004). Consider a function of two variables $u(s,t)$, the derivative of this function with respect to t when approximated by a forward difference scheme, is given by

$$\frac{\partial u(s,t)}{\partial t} = \frac{u(s,t+\Delta t) - u(s,t)}{\Delta t} + \mathcal{O}(\Delta t), \tag{3.1}$$

for a small Δt using a Taylor series expansion of u about (s,t). Using the same function, the backward difference scheme takes the form

$$\frac{\partial u(s,t)}{\partial t} = \frac{u(s,t) - u(s,t-\Delta t)}{\Delta t} + \mathcal{O}(\Delta t), \tag{3.2}$$

and expanding the terms in Equations 3.1 and 3.2, and subtracting one from the other, leads to the central difference approximation

$$\frac{\partial u(s,t)}{\partial t} = \frac{u(s,t+\Delta t) - u(s,t-\Delta t)}{2\Delta t} + \mathcal{O}((\Delta t)^2). \tag{3.3}$$

The main idea of the finite difference method is to discretise an original PDE to obtain a linear system of equations. The space axis (asset path) is divided into equally spaced nodes, spaced Δs apart and the same is done for the time axis, spaced Δt apart.

The application of these finite difference schemes gives rise to truncation errors, which are quoted by $\mathcal{O}(\Delta t)$ and $\mathcal{O}((\Delta t)^2)$ when using a particular finite difference approximation. Since Δt is a small quantity, then $\Delta t^2 < \Delta t$, meaning that the second order error represents a better approximation.

A direct Theta Method finite difference scheme is then derived for the Black-Scholes PDE for both the GBM and CEV diffusion processes. Consider the Black-Scholes PDE,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0, \tag{3.4}$$

where V is the option value, S is the underlying asset, σ is the volatility and r is the risk free interest rate.

The following section is adapted from McWalter (2016).

3.1.1 Theta method for GBM diffusion process

Although in theory the infinite domain $S \in [0,\infty)$, should be used as the bounds of the stock, the approach that will be followed below uses a truncated domain,

$S \in [S_{\min}, S_{\max}]$ for $0 \leq S_{\min} \leq S_{\max}$.

The Black-Scholes PDE above is transformed via a reversal of time, $\tau = T - t$, this enables the change of the terminal condition to be the initial condition. The time-reversed Black-Scholes equation then takes the form,

$$\frac{\partial U}{\partial \tau} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} - rS\frac{\partial U}{\partial S} + rU = 0. \tag{3.5}$$

A regular mesh is used to create the time and space (asset) variables: $\delta s = \frac{(S_{\max} - S_{\min})}{N}$, $\delta\tau = T/M$ and $\{(S_{\min} + n\delta s, m\delta\tau) : 0 \leq n \leq N, 0 \leq m \leq M\}$, where $M, N \in \mathbb{N}$.

Using Equations 3.1, 3.2 for the time derivatives and Equation 3.3 for the spatial derivatives, the following discretisation for the explicit and implicit schemes are derived respectively:

$$\begin{aligned} &\frac{U^n_{m+1} - U^n_m}{\delta\tau} - r(S_{\min} + n\delta s)\frac{U^{n+1}_m - U^{n-1}_m}{2\delta s} \\ &- \frac{1}{2}\sigma^2(S_{\min} + n\delta s)^2 \frac{U^{n+1}_m - 2U^n_m + U^{n-1}_m}{\delta s^2} + rU^n_m = 0, \end{aligned}$$

and

$$\begin{aligned} &\frac{U^n_{m+1} - U^n_m}{\delta\tau} - r(S_{\min} + n\delta s)\frac{U^{n+1}_{m+1} - U^{n-1}_{m+1}}{2\delta s} \\ &- \frac{1}{2}\sigma^2(S_{\min} + n\delta s)^2 \frac{U^{n+1}_{m+1} - 2U^n_{m+1} + U^{n-1}_{m+1}}{\delta s^2} + rU^n_{m+1} = 0. \end{aligned}$$

These equations may be written in matrix form as

$$\mathbf{U}_{m+1} = \mathbf{F}\mathbf{U}_m + \mathbf{b}_m, \tag{3.6}$$

$$\mathbf{G}\mathbf{U}_{m+1} = \mathbf{U}_m + \mathbf{b}_{m+1}, \tag{3.7}$$

for $0 \leq m \leq M-1$, where $\mathbf{U}_m, \mathbf{U}_{m+1} \in \mathbb{R}^{(N-1)}$ are solutions at times m and $m+1$, the term $\mathbf{b}_m$ is the vector that specifies the boundary conditions and $\mathbf{F}$ and $\mathbf{G}$ are $(N-1) \times (N-1)$ tridiagonal matrices.

The specifications of the matrices above are as follows:

$$\begin{aligned} &\mathbf{F} = (1 - r\delta\tau)\mathbf{I} + \frac{1}{2}r\delta\tau\mathbf{D}_1\mathbf{T}_1 + \frac{1}{2}\sigma^2\tau\mathbf{D}_2\mathbf{T}_2, \\ &\mathbf{G} = 2\mathbf{I} - \mathbf{F}, \\ &\mathbf{D}_1 = \mathrm{diag}(\frac{S_{\min}}{\delta s} + [1, 2, ..., N-1]), \\ &\mathbf{D}_2 = \mathbf{D}_1^2, \end{aligned}$$

the $(N-1)$ x $(N-1)$ tridiagonal matrices $\mathbf{T}_1, \mathbf{T}_2$ are defined by,

$$\mathbf{T}_1 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \ddots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & -1 & 0 \end{bmatrix} \mathbf{T}_2 = \begin{bmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & -2 \end{bmatrix}$$

and the vector of boundary conditions $\mathbf{b}_m$ which is of size $(N-1)$ is defined by,

$$\mathbf{b}_m = \begin{bmatrix} \frac{1}{2}\delta\tau(\frac{S_{\text{min}}}{\delta s}+1)(\sigma^2(\frac{S_{\text{min}}}{\delta s}+1)-r)U_m^0 \\ 0 \\ \vdots \\ 0 \\ \frac{1}{2}\delta\tau(\frac{S_{\text{max}}}{\delta s}-1)(\sigma^2(\frac{S_{\text{max}}}{\delta s}-1)+r)U_m^N \end{bmatrix}.$$

These boundary conditions are specified as $U_m^0 = U^0(S_{\text{min}}, m\delta\tau)$ and $U_m^N = U^\infty(S_{\text{max}}, m\delta\tau)$.

To solve the system of equations the initial condition is specified as $U_0^n = U_0(S_{\text{min}} + n\delta s) = V_T$, which is the terminal payoff of the option, $C(S,T) = \max(S-K, 0)$, and $P(S,T) = \max(K-S, 0)$. The boundary conditions, U^0, U^∞ are also required along with the option related parameters σ(volatility), r(interest rate), K(strike) and the mesh related parameters $S_{\text{min}}, S_{\text{max}}, N, M$.

To derive the Theta Method, Equation 3.6 is multiplied by $(1-\theta)$ and Equation 3.7 by θ and then the two equations are summed together:

$$\theta\mathbf{G} + (1-\theta)\mathbf{U}_{m+1} = ((1-\theta)\mathbf{F} + \theta\mathbf{I})\mathbf{U}_m + (1-\theta)\mathbf{b}_m + \theta\mathbf{b}_{m+1},$$

then solving for $\mathbf{U}_{m+1}$ gives the solution at time step $m+1$,

$$\mathbf{U}_{m+1} = (\theta\mathbf{G} + (1-\theta))^{-1}[((1-\theta)\mathbf{F} + \theta\mathbf{I})\mathbf{U}_m + (1-\theta)\mathbf{b}_m + \theta\mathbf{b}_{m+1}]. \tag{3.8}$$

The above method then needs to be adapted for the Bermudan put option given that the option has specific exercise dates. If the current date, t_{m+1}, coincides with the date that the option can be exercised, then the price of the Bermudan option is given by,

$$\begin{aligned} \mathbf{U}_{m+1} = \max((\theta\mathbf{G} + (1-\theta))^{-1}[((1-\theta)\mathbf{F} + \theta\mathbf{I})\mathbf{U}_m \\ + (1-\theta)\mathbf{b}_m + \theta\mathbf{b}_{m+1}], \max(K-S, 0)). \end{aligned} \tag{3.9}$$

If the exercise price of the option $\max(K-S, 0)$, is greater than the continuation value, $\mathbf{U}_{m+1}$, then the option is exercised. Otherwise, if the date does not coincide with the early exercise dates, the option price is simply given by Equation 3.8.

This scheme can be shown to be unconditionally stable (McWalter, 2016) for $\frac{1}{2} \leq \theta \leq 1$. The parameter choice $\theta = 1$ represents the fully implicit case while $\theta = 0$ represents the fully explicit case and $\theta = \frac{1}{2}$ represents the Crank-Nicolson scheme.

3.1.2 Theta method for the CEV diffusion process

The formulation of this method is similar to the above formulation of the Theta Method for the GBM diffusion process, only certain relevant terms are changed.

Again, consider the time-reversed Black-Scholes PDE but under the CEV diffusion process:

$$\frac{\partial U}{\partial \tau} - \frac{1}{2}\sigma^2(S, T-\tau)S^2\frac{\partial^2 U}{\partial S^2} - rS\frac{\partial U}{\partial S} + rU = 0, \tag{3.10}$$

slight variations to the formulation above are made, where $\sigma(S, T-\tau)$ is now the instantaneous volatility of the percentage price change which is equal to $\sigma_{LN}(S_t) = \sigma S_t^{\alpha-1}$. The volatility is now inversely proportional to the stock price as was defined in the second chapter. A new matrix is defined,

$$\mathbf{\Sigma}_m = \begin{bmatrix} \sigma^2(t_m, S_{\text{min}} + \delta s) & 0 & \dots & 0 \\ 0 & \sigma^2(t_m, S_{\text{min}} + 2\delta s) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma^2(t_m, S_{\text{min}} + (N-1)\delta s) \end{bmatrix},$$

with $t_m = T - m\delta\tau$. Then define two new matrix formulations:

$$\mathbf{F}_m = (1 - r\delta\tau)\mathbf{I} + \frac{1}{2}r\delta\tau\mathbf{D}_1\mathbf{T}_1 + \frac{1}{2}\Sigma_m\tau\mathbf{D}_2\mathbf{T}_2, \tag{3.11}$$

$$\mathbf{G}_{m+1} = (1 + r\delta\tau)\mathbf{I} - \frac{1}{2}r\delta\tau\mathbf{D}_1\mathbf{T}_1 - \frac{1}{2}\Sigma_{m+1}\tau\mathbf{D}_2\mathbf{T}_2, \tag{3.12}$$

different to the above theta scheme for the GBM case, where $\mathbf{F}_m$ and $\mathbf{G}_{m+1}$ are tridiagonal matrices as was defined earlier at times m and $m+1$. Then new boundary conditions are defined,

$$\mathbf{b}_m = \begin{bmatrix} \frac{1}{2}\delta\tau(\frac{S_{\text{min}}}{\delta s}+1)(\sigma^2(t_m, S_{\text{min}}+\delta s)(\frac{S_{\text{min}}}{\delta s}+1)-r)U_m^0 \\ 0 \\ \vdots \\ 0 \\ \frac{1}{2}\delta\tau(\frac{S_{\text{max}}}{\delta s}-1)(\sigma^2(t_m, S_{\text{max}}-\delta s)(\frac{S_{\text{max}}}{\delta s}-1)+r)U_m^N \end{bmatrix},$$

but with $U_m^0 = U^0(S_{\text{min}}, m\delta\tau)$ and $U_m^N = U^\infty(S_{\text{max}}, m\delta\tau)$ as described before.

Using the above matrices and the correct initial and boundary conditions, the solution to Equation 3.10 is given by Equation 3.8 with the formulation of $\mathbf{F}$ changing to $\mathbf{F}_m$ and the formulation of $\mathbf{G}$ changing to $\mathbf{G}_{m+1}$.

For the Bermudan put option, if the current date and the early exercise date do not coincide, the price is given by Equation 3.8, with the formulation of $\mathbf{F}$ changed to $\mathbf{F}_m$ and the formulation of $\mathbf{G}$ changed to $\mathbf{G}_{m+1}$. If the early exercise date and the current date do coincide, the solution is given by whichever is greater: the early exercise price, $\max(K-S,0)$ or the continuation value, $\mathbf{U}_{m+1}$. Thus the Bermudan put option price is given by Equation 3.9 with $\mathbf{F}$ and $\mathbf{G}$ changed as above.

3.2 Least squares Monte Carlo method

In this section, the method of Longstaff and Schwartz (2001) to price a Bermudan put option is reviewed. The least squares Monte Carlo method is a simulation method that combines a Monte Carlo simulation with a least squares regression. The core to this method is to determine the conditional expected option value (continuation value) by generating sample paths and performing a regression analysis on the resulting option values. The value for this regression then represents the value of continuing to hold the option (Longstaff and Schwartz, 2001), which can be compared to the early exercise value, at that time for a particular state.

To explain the intuition in further detail, recall that a Bermudan option can be exercised at a restricted number of dates, the challenge then is to determine when the holder of a Bermudan option chooses to stop. The procedure is to find the optimal exercise rule and compute the expected discounted payoff using the optimal exercise rule. This optimal exercise rule is based on price up to the present moment.

The idea is to work backward in time, using a backward dynamic programming approach, consider the set of discrete times $0 < t_1 < t_2 < \cdots < t_N = T$, to value a Bermudan option. Let E represent the payoff at time t, $E(S) = (K - S)^+$. Define the value of the option to be U.

Then the construction of $U(t)$ using backward induction is,

$$U_N(s) = E(s)$$
$$U_{i-1}(s) = \max\{E_{i-1}(s), \mathbb{E}_{\mathbb{Q}}[U_i(S_i)e^{-r\Delta t_i}|S_{i-1} = s]\}$$

for $N \geq i \geq 1$, where $\Delta t_i = t_i - t_{i-1}$. $U(t)$ is called the Snell envelope. The expectation is taken under the risk neutral measure, where the stock evolves with a drift equal to the risk free rate.

Thus at any exercise time, one needs to compare the payoff from immediate exercise with the continuation value, $\mathbb{E}_{\mathbb{Q}}[U_i(S_i)e^{-r\Delta t_i}|S_{i-1} = s]$, and exercise if the immediate payoff is greater. The continuation value is the discounted option value that has been regressed, it is interpreted as the expected value of the option conditioned on the option not having been exercised before t_{i-1} and the share price having the value $S_{i-1} = s$ (McWalter, 2016). Estimating this conditional expectation is the challenge in valuing options with an early exercise feature.

The key insight of Longstaff and Schwartz (2001) was to identify that the conditional expectation can be estimated using regression on the cross-sectional information at a particular time step in a Monte Carlo simulation. This can be achieved by assuming the continuation value is specified by a suitable parametric function $f(\hat{\beta}_{i-1}, x)$, where $\hat{\beta}_{i-1}$ is found by regressing realised payoffs at each exercise time against asset prices. Using this estimate of the conditional expectation, it is possible to decide pathwise whether it is optimal to exercise early or if continuation should occur.

The form of the continuation function will be specified by

$$f(\hat{\beta}, x) = \sum_{r=0}^{R} \hat{\beta}_r \phi_r(x)$$

where the $\phi_r(x)$ terms the Laguerre polynomials which have been chosen as the basis functions. The choice of basis function has a significant effect in accurately describing the continuation value, while the approximation error also depends on the choice of regressor. Below are the Laguerre polynomials described above,

$$\phi_0(x) = 1, \qquad \phi_1(x) = 1 - x, \qquad \phi_2(x) = 1 - 2x + \frac{x^2}{2}, \qquad \phi_r(x) = \frac{e^x}{r!}\frac{d^r}{dx^r}(x^r e^{-x}).$$

There are many other possible choices for the basis functions namely, the Hermite, Legendre, Chebyshev, Gegenbauer and the Jacobi polynomials.

The best way to describe this method is by describing the least squares Monte Carlo algorithm. Below are the steps for the pricing of a Bermudan put option. Unlike the method in the previous section, this method works only for pricing an option with an early exercise feature.

3.2.1 Algorithm

Above, a high level theoretical understanding is given, and the algorithm below is presented. To begin, define all option related parameters:

- S_0 (initial stock price),
- T (option maturity),
- K (strike),
- r (risk free interest rate),
- α (elasticity parameter, in the case of the CEV diffusion process),
- σ (volatility),
- n, number of paths,
- N, number of steps (exercise dates),
- $\phi_0 \ldots \phi_R$, basis functions.

To begin,

1. Draw N x n, normally distributed numbers Z. Using the option related parameters, generate stock price paths using the Euler-Maruyama scheme described below with N steps and n paths,
$$S_{t_i} = S_{t_{i-1}} + rS_{t_{i-1}}\Delta t + \sigma S_{t_{i-1}}^{\alpha}\sqrt{\Delta t}Z_n \tag{3.13}$$
for $1 \leq i \leq N$, where $\Delta t = \frac{T}{N}$.
2. Compute the terminal payoff, $V_N = \max(K - S_T, 0)$ for each path.

3. For $i = N, N-1, \ldots, 2$ repeat steps 4 - 7.

4. Compute the realised continuation value, $V_{i-1} = e^{-r\Delta t_i} V_i$ for each path,

5. Let p be the paths for which early exercise is greater than zero.

6. Let X be the vector of the stock prices for these paths, p, and Y be the associated realised continuation values. Then perform a least squares regression on Y and $f(\hat{\beta}, X)$ to produce an estimate of $\hat{\beta}$.

7. For paths where early exercise is greater than $f(\hat{\beta}, X)$, set V_{i-1} equal to the early exercise values.

8. Then the value of the option at the initial time is given by $V_0 = \mathbb{E}[e^{-r\Delta t} V_1]$.

3.3 Recursive marginal quantization method

In this chapter, a new approach that was proposed by Pagès and Sagna (2015) as an efficient numerical method for evaluating functionals of solutions of stochastic differential equations is described. This work has been extended by McWalter *et al.* (2017), to show that it is possible to perform recursive marginal quantization for two higher-order schemes: the Milstein scheme and a simplified weak order 2.0 scheme. There are two parts to this method, the first is vector quantization and the second is recursive marginal quantization.

It is important to take note that, in order to solve this problem, many equations have to be derived. The material presented here is merely a summary of the entire process (see McWalter *et al.* (2017) for detailed derivations).

3.3.1 Vector quantization

Vector quantization is a technique from signal processing that allows the approximation of a continuous probability density function by a discrete probability mass function. Vector quantization is known as a lossy data compression technique (McWalter *et al.*, 2017) because after compression, it is no longer possible to reconstruct the original exactly.

The problem that vector quantization addresses is, how does one optimally approximate the continuous distribution associated with the random variable, X, in a least squares sense, using a discrete random variable, $\hat{X} : \Omega \to \Gamma$, where Γ is a finite set of elements in $\mathbb{R}$ (McWalter *et al.*, 2017)?

The technique concerns itself with the best approximation of a one dimensional probability distribution. Consider the approximation of X, given by $\hat{X} = \text{Proj}_\Gamma(X)$, a discrete random vector that consists of a finite number of points defined as the nearest-neighbour projection of X onto $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_N\}$. Here Γ is known as a quantizer with its elements known as codewords. The nearest neighbour projection operator, $\text{Proj}_\Gamma : \mathbb{R} \to \Gamma$, is defined as,

$$\text{Proj}_\Gamma(X) = \{\gamma^i \in \Gamma | ||X - \gamma^i|| \leq ||X - \gamma^j||\},$$

for all $j = 1, \ldots, N$; where equality holds for $i < j$. Intuitively, $\hat{X} = \text{Proj}_\Gamma(X)$ of which the $\text{Proj}_\Gamma(X)$ is to be γ^i, which are the elements of the quantizer known as the codewords. The definition further states that i is the smallest k for which $||X - \gamma^k|| = \min^j ||X - \gamma^j||$.

Associated with the quantizer, the regions $R_i(\Gamma) \subset \mathbb{R}$ are the subset of values of X that are mapped to each codeword γ^i:

$$R_i(\Gamma) = \{x \in \mathbb{R} | \text{Proj}_\Gamma(x) = \gamma^i\}.$$

These are known as Voronoi regions.

Given that this is a one dimensional problem, the regions may be defined explicitly as $R_i = \{x | r^{i-} < x \leq r^{i+}\}$, with,

$$r^{i-} = \frac{\gamma^{i-1} - \gamma^i}{2} \qquad r^{i+} = \frac{\gamma^i - \gamma^{i+1}}{2},$$

for $1 \leq i \leq N$, where by definition $r^1 = -\infty$, and $r^N = \infty$. If the distribution is not defined on the whole real line then, r^1, r^N will reflect the appropriate support of the distribution function.

The problem that one wishes to solve is to find a quantizer, Γ, such that $\hat{X} = \text{Proj}_\Gamma(X)$ best approximates X. To accomplish this, one minimizes the distortion function given by,

$$D(\Gamma) = \mathbb{E}[||X - \hat{X}||^2].$$

A way to solve such a system is to represent the quantizer by a column vector $\mathbf{\Gamma}$, derive the gradient $\nabla D(\mathbf{\Gamma})$ and the Hessian, $\nabla^2 D(\mathbf{\Gamma})$, and then use a Newton-Rhapson iteration (defining the set Γ^n with the vector $\mathbf{\Gamma}^n$),

$$\mathbf{\Gamma}^{n+1} = \mathbf{\Gamma}^n - [\nabla^2 D(\mathbf{\Gamma}^n)]^{-1} \nabla D(\mathbf{\Gamma}^n). \tag{3.14}$$

Before the Newton-Rhapson iteration can be solved a few terms need to be defined.

Let F_X, f_X be the CDF and PDF of X, then define the $p-th$ lower partial expectation to be,

$$M_X^p(x) = \mathbb{E}[X\mathbb{I}_{\{X<x\}}]. \tag{3.15}$$

Now that the relevent terms have been defined, an efficient implementation of the Newton-Rhapson iteration (as per McWalter *et al.* (2017)) above is presented below.

Let

$$[\mathbf{\Gamma}]_i = \gamma_i, \qquad [\mathbf{M}]_i = M_X(r^{i+}) - M_X(r^{i-}), \qquad 1 \leq i \leq N,$$

$$[\mathbf{f}]_i = f_X(r^{i+}), \qquad [\mathbf{\Delta\Gamma}]_i = \gamma^{i+1} - \gamma^i \qquad 1 \leq i \leq N-1.$$

Then define the row vector of probabilities, $\mathbf{p}$, as follows

$$[\mathbf{p}]_i = F_X(r^{i+}) - F_X(r^{i-}), \qquad \text{for} \quad 1 \leq i \leq N.$$

Using the above vectors, the gradient of the distortion function is then,

$$\nabla \mathbf{D}(\mathbf{\Gamma}) = 2\mathbf{\Gamma} \circ \mathbf{p}^\intercal - \mathbf{2M},$$

where $\circ$ is the elementwise Hadamard product. Then the entries of the off diagonals and main diagonal tridiagonal Hessian, $\nabla^2 D(\mathbf{\Gamma})$ are defined,

$$\mathbf{h_{off}} = -\frac{1}{2}[\mathbf{f} \circ \mathbf{\Delta\Gamma}]^\intercal,$$
$$\mathbf{h_{main}} = 2\mathbf{p} + [\mathbf{h_{off}}|0] + [0|\mathbf{h_{off}}],$$

respectively, where the $\mathbf{h_{off}}$ vector is appended and prepended with zero. After setting up the above vectors, the Newton-Rhapson iteration can be computed. The figures below show vector quantization applied to a Gaussian distribution and a non-central chi-squared distribution with one degree of freedom.

The Standard Normal Distribution

If X is a standard normal random variable then,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} = \phi(x),$$
$$F_X(x) = \frac{1}{2}[1 + \text{erf}(\frac{x}{\sqrt{2}})] = \Phi(x),$$
$$M_X(x) = -\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} = -\phi(x),$$

and a good guess for the initial quantizer, $\mathbf{\Gamma}^0$ is

$$\gamma^n = \frac{5.5n}{N+1} - 2.75 \qquad \text{for } 1 \leq n \leq N,$$

Figure 3.1 shows the quantizer of cardinality $N = 50$, after computing 50 Newton-Rhapson iterations.

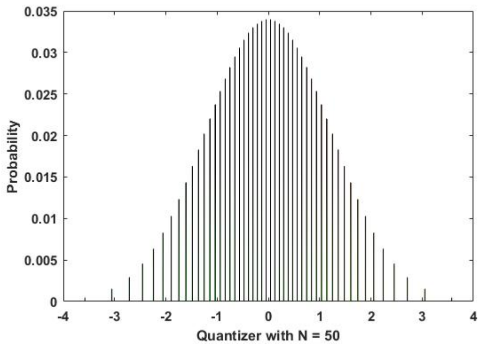

Fig. 3.1: Vector quantization of the Standard Normal distribution

The Noncentral Chi-squared Distribution

If $X \sim \chi'^2(1, \lambda)$ is a noncentral chi-squared distributed random variable with one degree of freedom and noncentrality parameter $\lambda = \mu^2$ then,

$$f_X(x) = \frac{1}{2\sqrt{x}}(\phi(x^+) + \phi(x^-)),$$
$$F_X(x) = \Phi(x^+) - \Phi(x^-),$$
$$M_X^1(x) = (\lambda + 1)(\Phi(x^+) - \Phi(x^-)) + \phi(x^+)x^- - \phi(x^-)x^+,$$

where

$$x^{\pm} = \pm\sqrt{x} - \sqrt{\lambda}.$$

A good initial guess for the quantizer $\mathbf{\Gamma}^0$ is

$$\begin{cases} \gamma^n = (\frac{(3+\sqrt{\lambda})n}{N})^2 & \text{for } \sqrt{\lambda} < 2.5 \\ \gamma^n = (\frac{5n}{N+1} - 2.5 + \sqrt{\lambda})^2 & \text{for } \sqrt{\lambda} \geq 2.5, \end{cases}$$

for $1 \leq n \leq N$. Figure 3.2 shows the quantizer of cardinality $N = 50$ for the noncentral chi-squared distribution with one degree of freedom and noncentrality parameter, $\lambda = 5$. Similar to the vector quantization of the standard Normal distribution, 50 Newton-Rhapson iterations were computed.

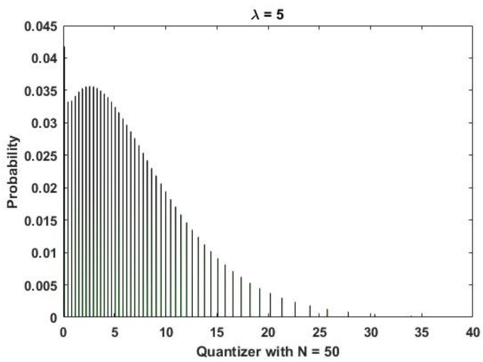

Fig. 3.2: Vector quantization of the noncentral Chi-squared distribution, $\lambda = 5$.

3.3.2 Recursive marginal quantization

Consider a new approach to quantize the marginals of a stochastic diffusion process at discrete times. This method was proposed by Pagès and Sagna (2015) for the quantization of the marginals of a discrete Euler diffusion process and has been extended by McWalter *et al.* (2017) for higher order schemes. The analysis below is restricted to the one dimensional setting.

Let $(X_t)_{t\geq 0}$ be a stochastic process, defined by

$$dX_t = a(t, X_t)\, dt + b(t, X_t)\, dW_t, \qquad X_0 = x_0 \in \mathbb{R}, \tag{3.16}$$

where W_t is the standard Brownian motion, with a and b being sufficiently smooth and regular functions to ensure the existence of a weak solution. The problem then is, how to approximate $X_{t_k} : \Omega \to \mathbb{R}$ optimally for some time discretisation, $t_k \in [0, T]$, when the distribution of X_{t_k} is unknown?

The Euler-Maruyama scheme

Consider the Euler-Maruyama scheme of the process, X_{t_k}, with an initial condition of $\bar{X}_0 = X_0$:

$$\bar{X}_{t_{k+1}} = \bar{X}_{t_k} + a(t_k, \bar{X}_{t_k})\Delta t + b(t_k, \bar{X}_{t_k})\sqrt{\Delta t} Z_{t_{k+1}} =: U(\bar{X}_{t_k}, Z_{t_{k+1}})$$

for $0 \leq t_k \leq K$, where $\Delta t = \frac{T}{K}$ and $Z_{t_{k+1}}$ are independent $\mathcal{N}(0,1)$ random variables. To alleviate notation, $\bar{X}_{t_k} = \bar{X}_k$ and $Z_{t_{k+1}} = Z_{k+1}$.

Given that $\bar{X}_1$ is Gaussian distributed, vector quantization can be applied to optimally approximate $\mathbf{\Gamma}_1$, a first step in the recursive marginal quantization method. The distortion function associated with the quantizer of $\bar{X}_{k+1}$ may be written as

$$\bar{D}(\Gamma_{k+1}) = \mathbb{E}\big[||\bar{X}_{k+1} - \hat{X}_{k+1}||^2\big] = \mathbb{E}\big[\mathbb{E}\big[||\mathcal{U}(\bar{X}_k, Z_{k+1}) - \hat{X}_{k+1}||^2|\bar{X}_k\big]\big].$$

The distribution of $\bar{X}_k$ is, however, unknown. Pagès and Sagna (2015) showed that the recursive marginal quantization process converges, if one uses the previously quantized distribution of $\bar{X}_k$, instead of the continuous distribution of $\bar{X}_k$ to find the successive marginal distributions of $\bar{X}_{k+1}$. Furthermore, the approximate value for the distortion may then be written as the sum over the codewords in quantizer $\mathbf{\Gamma}_k$ and their associated probabilities,

$$\bar{D}(\Gamma_{k+1}) \approx D(\Gamma_{k+1}) = \sum_{i=1}^{N_k} \mathbb{E}\big[||\mathcal{U}(\gamma^i, Z_{k+1}) - \hat{X}_{k+1}||^2\big]\mathbb{P}(\hat{X}_k = \gamma_k^i).$$

The N_k represents the cardinality of the quantizer Γ_k at time step k. Given the quantizer expressed as a column vector $\mathbf{\Gamma}_k$ and the associated probabilities $\mathbb{P}(\hat{X}_k = \gamma_k^i)$, for $1 \leq i \leq N_k$, the Newton-Rhapson iteration for the quantizer Γ_{k+1} at time step t_{k+1} is then

$$\mathbf{\Gamma}_{k+1}^{n+1} = \mathbf{\Gamma}_{k+1}^{n} - \big[\nabla^2 D(\mathbf{\Gamma}_{k+1}^{n})\big]^{-1} \nabla D(\mathbf{\Gamma}_{k+1}^{n}), \tag{3.17}$$

for $0 \leq n < n_{max}$. The initial guess is given by $\mathbf{\Gamma}_{k+1}^0 = \mathbf{\Gamma}_k$.

Before the Newton-Rhapson iterations can be computed, some more quantities need to be defined. The update for the Euler-Maruyama scheme may be written as

$$\mathcal{U}(\gamma^i, Z_{k+1}) = m_k^i Z_{k+1}^i + c_k^i, \tag{3.18}$$

where

$$m_k^i = b(t_k, \gamma_k^i)\sqrt{\Delta t} \qquad c_k^i = \gamma_k^i + a(t_k, \gamma_k^i)\sqrt{\Delta t},$$

with $Z^i_{k+1} \sim \mathcal{N}(0,1)$. The regions, $R_j(\Gamma_{k+1})$ associated with the quantizer where $U^i_{k+1} \in R_j(\Gamma_{k+1})$, are defined as

$$r^{j-}_{k+1} < U^i_{k+1} \leq r^{j+}_{k+1} \qquad r^{j\pm}_{k+1} = \frac{1}{2}(\gamma^{j\pm 1}_{k+1} - \gamma^j_{k+1}),$$

for $1 \leq j \leq N_{k+1}$, where $r^{1-}_{k+1} = -\infty$ and $r^{N_{k+1}+}_{k+1} = \infty$. The conditionally normalized boundaries are then defined by

$$r^{i,j\pm}_{k+1} = \frac{r^{j\pm}_{k+1} - c^i_k}{m^i_k}.$$

As before, an efficient implementation of the Newton-Rhapson iteration is presented below (See McWalter *et al.* (2017) for details).

Let,

$$[\mathbf{m}_k]_i = m^i_k, \qquad [\mathbf{c}_k]_i = c^i_k, \qquad 1 \leq i \leq N_k,$$

which represents the time-indexed vectors, and

$$[\boldsymbol{\Delta}\boldsymbol{\Gamma}_{k+1}]_i = \gamma^{i+1}_{k+1} - \gamma^i_{k+1}, \qquad 1 \leq i \leq N_{k+1} - 1.$$

Then define the row of probabilities

$$[\mathbf{p}]_k = [\mathbb{P}(\hat{X}_k = \gamma^1_k), \ldots, \mathbb{P}(\hat{X}_k = \gamma^{N_k}_k)], \tag{3.19}$$

and define a row of ones of length d denoted by $\mathbf{j}_d$. Note that $\boldsymbol{\Delta}\boldsymbol{\Gamma}_{k+1}$ must be computed before each Newton-Rhapson iteration, while the other vectors are computed once per time-step.

Define three matrices, which are computed in terms of the new estimate of Γ_{k+1}

$$[\mathbf{P}_{k+1}]_{i,j} = \text{sgn}(m^i_k)[F_{Z^i_{k+1}}(r^{i,j+}_{k+1}) - F_{Z^i_{k+1}}(r^{i,j-}_{k+1})], \tag{3.20}$$

which is an N_k x N_{k+1} matrix of transition probabilities. Then set the lower partial moment values to

$$[\mathbf{M}_{k+1}]_{i,j} = M^1_{Z^i_{k+1}}(r^{i,j+}_{k+1}) - M^1_{Z^i_{k+1}}(r^{i,j-}_{k+1}), \tag{3.21}$$

which is the same size as the matrix of transition probabilities. Define the matrix of density values at the positive region boundaries by

$$[\mathbf{f}_{k+1}]_{i,j} = f_{Z^i_{k+1}}(r^{i,j+}_{k+1}), \tag{3.22}$$

an $N_k \times N_{k+1} - 1$ matrix. Then define the gradient of the distortion function as

$$\nabla D(\mathbf{\Gamma}_{k+1})^\intercal = 2\mathbf{p}_k(((\mathbf{\Gamma}_{k+1}\mathbf{j}_{N_k})^\intercal - \mathbf{c}_k\mathbf{j}_{N_{k+1}}) \circ \mathbf{P}_{k+1} - (|m_k|\mathbf{j}_{N_{k+1}}) \circ \mathbf{M}_{k+1}), \quad (3.23)$$

where $\circ$ is the Hadamard product. Then define the entries of the Hessian, $\nabla^2 D(\mathbf{\Gamma})$,

$$\begin{aligned} \mathbf{h_{off}} &= -\frac{1}{2}\mathbf{p}_k((|\mathbf{m}_k|^{\circ-1}\mathbf{j}_{N_{k+1}-1}) \circ \mathbf{f}_{k+1} \circ (\Delta\mathbf{\Gamma}_{k+1}\mathbf{j}_{N_k})^\intercal), \\ \mathbf{h_{main}} &= 2\mathbf{p}_k\mathbf{P}_{k+1} + [\mathbf{h_{off}}|0] + [0|\mathbf{h_{off}}], \end{aligned} \quad (3.24)$$

where $\circ - 1$ is the element-wise inverse.
After setting up the above vectors and matrices, the Newton-Rhapson iteration can be computed. The probabilities associated with the final quantizer are

$$\mathbf{p}_{k+1} = \mathbf{p}_k\mathbf{P}_{k+1}, \quad (3.25)$$

as mentioned before, $\mathbf{P}_{k+1}$ must be computed for each $\mathbf{\Gamma}_{k+1}$. Figure 3.3 below shows the evolution of the quantizers, for the CEV and GBM diffusion processes, where the stock price evolves with the following dynamics

$$dS_t = \mu S_t\,dt + \sigma S_t\,dW_t,$$

for the GBM case, using parameters $S_0 = 100$, $\sigma = 0.3$, $r = 0.05$, $K = 12$, strike $= 100$, $T = 1$,

and

$$dS_t = \mu S_t\,dt + \sigma S_t^\alpha\,dW_t,$$

for the CEV case, using parameters $S_0 = 100$, $\sigma_{LN} = 0.3$, $r = 0.05$, $K = 12$, strike $= 100$, $\alpha = 0.7$, $\sigma = \sigma_{LN}S_0^{1-\alpha}$, $T = 1$.
In these plots, the different colours of the quantizer indicate the associated time step, the blue line in the middle represents the initial time and the lines extending beyond the blue line are closer to the final time.

McWalter *et al.* (2017) showed that by generalizing the RMQ procedure, it is possible to perform recursive marginal quantization on higher-order schemes, this being the Milstein scheme and the weak order 2.0 Taylor scheme.

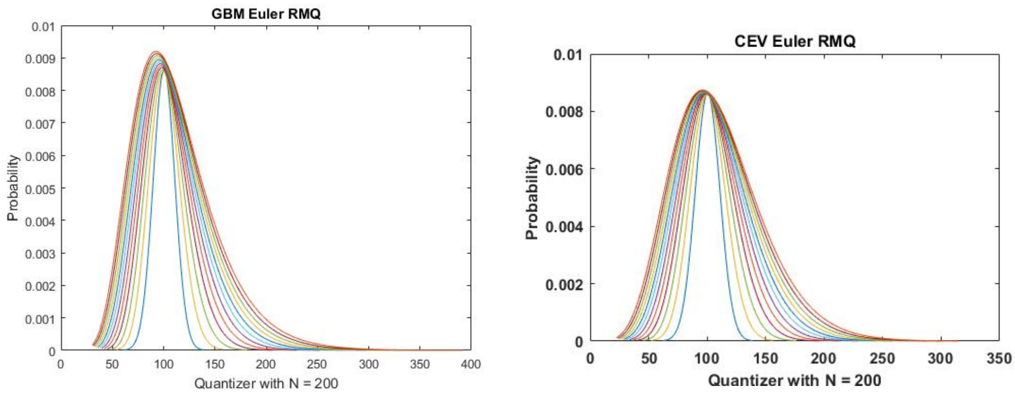

Fig. 3.3: GBM and CEV process for the Euler scheme

The Milstein scheme

Consider the Milstein scheme for the Equation 3.16, with initial condition $\bar{X}_0 = x_0$:

$$\begin{aligned}\bar{X}_{k+1} = \bar{X}_k &+ (a(\bar{X}_k) - \frac{1}{2}b(\bar{X}_k)b^{'}(\bar{X}_k))\Delta t \\ &- \frac{1}{2}b(\bar{X}_k)b^{'}(\bar{X}_k)^{-1} + \frac{1}{2}b(\bar{X}_k)b^{'}(\bar{X}_k)\Delta t(Z_{k+1} + (\sqrt{\Delta t}b^{'}(\bar{X}_k))^{-1})^2,\end{aligned}$$

for $0 \leq k < K$, where $\Delta t = \frac{T}{K}$ and Z_{k+1} is independent $\mathcal{N}(0, 1)$.

The update for this scheme is then,

$$\mathcal{U}(\gamma^i, Z_{k+1}) = m_k^i Z_{k+1}^i + c_k^i,$$

the variable Z_{k+1}^i is now noncentral chi-squared distributed with one degree of freedom and non-centrality parameter, $\lambda_{k+1}^i = (\sqrt{\Delta t}b^{'}(\gamma_k^i))^{-2}$, the distribution of the random variable now depends on the codeword γ_k^i as $Z_{k+1}^i \sim \chi^{'2}(1, \lambda_{k+1}^i)$. Here

$$m_k^i = \frac{1}{2}b(\gamma_k^i)b^{'}(\gamma_k^i)\Delta t,$$

and

$$c_k^i = \gamma_k^i + (a(\gamma_k^i) - \frac{1}{2}b(\gamma_k^i)b^{'}(\gamma_k^i)) - \frac{1}{2}b(\gamma_k^i)b^{'}(\gamma_k^i)^{-1}.$$

Figure 3.4 below shows the evolution of the quantizer for both the CEV and GBM diffusion processes for the Milstein scheme. The parameters are the same as the Euler case.

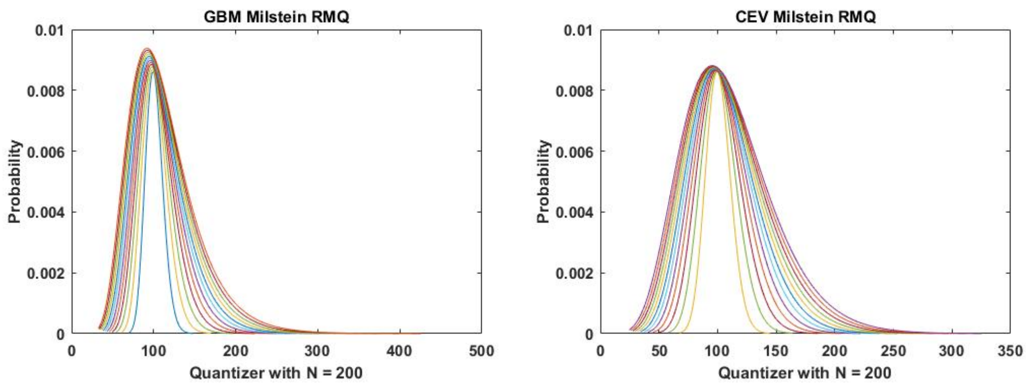

Fig. 3.4: GBM and CEV process for the Milstein scheme

A weak order 2.0 Taylor scheme

Consider the simplified weak order 2.0 scheme for Equation 3.16, with initial condition $\bar{X}_0 = x_0$:

$$\begin{aligned}\bar{X}_{k+1} = \bar{X}_k + a(\bar{X}_k)\Delta t + b(\bar{X}_k)\sqrt{\Delta t}Z_{k+1} + \frac{1}{2}b(\bar{X}_k)b^{'}(\bar{X}_k)\Delta t(Z_{k+1}^2 - 1)\\ + \frac{1}{2}(a^{'}(\bar{X}_k)b(\bar{X}_k) + a(\bar{X}_k)b^{'}(\bar{X}_k) + \frac{1}{2}b^{''}(\bar{X}_k)b^2(\bar{X}_k))(\Delta t)^{\frac{3}{2}}Z_{k+1}\\ + \frac{1}{2}(a(\bar{X}_k)a^{'}(\bar{X}_k) + \frac{1}{2}a^{''}(\bar{X}_k)b^2(\bar{X}_k))(\Delta t)^2,\end{aligned}$$

for $0 \le k < K$, where $\Delta t = \frac{T}{K}$ and Z_{k+1} is independent $\mathcal{N}(0,1)$.

The update for this scheme is then

$$\mathcal{U}(\gamma^i, Z_{k+1}) = m_k^i Z_{k+1}^i + c_k^i, \tag{3.26}$$

the variable Z_{k+1}^i is noncentral chi-squared distributed with one degree of freedom and non-centrality parameter,

$$\lambda_{k+1}^i = \left(\frac{b(\gamma_k^i) + \frac{1}{2}(a^{'}(\gamma_k^i)b(\gamma_k^i) + a(\gamma_k^i)b^{'}(\gamma_k^i) + \frac{1}{2}b^{''}(\gamma_k^i)b^2(\gamma_k^i))\Delta t}{b(\gamma_k^i)b^{'}(\gamma_k^i)\sqrt{\Delta t}}\right)^2,$$

the distribution of the random variable depends on the codeword γ_k^i as $Z_{k+1}^i \sim \chi^{'2}(1, \lambda_{k+1}^i)$. Here,

$$m_k^i = \frac{1}{2}b(\gamma_k^i)b^{'}(\gamma_k^i)\Delta t$$

and,

$$\begin{aligned}c_k^i = \gamma_k^i + (a(\gamma_k^i) - \frac{1}{2}b(\gamma_k^i)b^{'}(\gamma_k^i))\Delta t + \frac{1}{2}(a(\gamma_k^i)a^{'}(\gamma_k^i) + \frac{1}{2}a^{''}(\gamma_k^i)b^2(\gamma_k^i))(\Delta t)^2\\ - \frac{(b(\gamma_k^i) + \frac{1}{2}(a^{'}(\gamma_k^i)b(\gamma_k^i) + a(\gamma_k^i)b^{'}(\gamma_k^i) + \frac{1}{2}b^{''}(\gamma_k^i)b^2(\gamma_k^i))\Delta t)^2}{2b(\gamma_k^i)b^{'}(\gamma_k^i)}.\end{aligned}$$

Figure 3.5 below shows the evolution of the quantizer for both the CEV and GBM diffusion processes for the weak order 2.0 scheme. The parameters have been kept constant as in the Euler case.

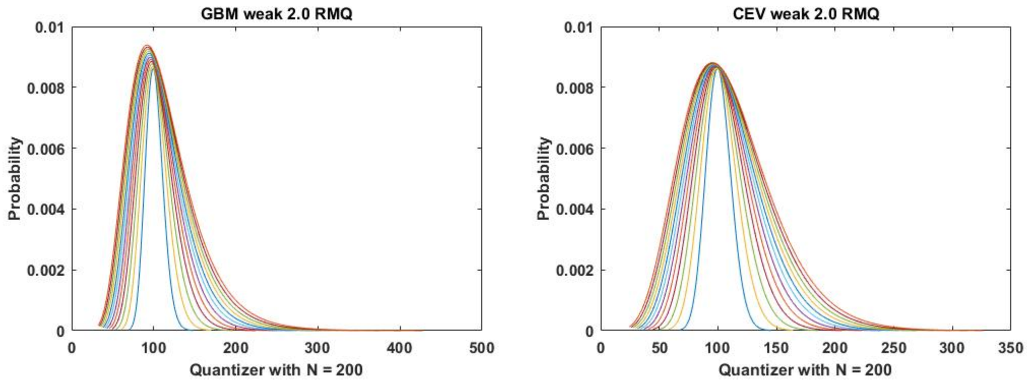

Fig. 3.5: GBM and CEV process for the weak order 2.0 Taylor scheme

3.4 Results

In this section, a European put option and a Bermudan put option are priced using the finite difference method, least squares Monte Carlo method and recursive marginal quantization method under the geometric Brownian motion (GBM) and constant elasticity of variance (CEV) processes. All three methods are then compared for computational efficiency and accuracy. Below is a table of general parameters that are used by all three methods

Parameters		
General	GBM	CEV
S_0	100	100
r	0.05	0.05
T	1	1
K	100	100
σ_{LN}	N/A	0.3
σ	0.3	$\sigma_{LN} S_0^{1-\alpha}$

Tab. 3.1: General price parameters used for all three methods.

In the case of the Bermudan put option price, the number of permissible option pricing dates is 12 (monthly exercise), this means that for RMQ, $K = 12$.

3.4.1 European put option price

Here, the European put option is priced using the following methods, finite difference method, the RMQ Euler-Maruyama scheme, the RMQ Milstein scheme, the RMQ weak order 2.0 Taylor scheme.

Figure 3.6 shows the plot of accuracy as a function of moneyness, $\frac{K}{S_0}$, where strike is varied. The accuracy was determined by taking the absolute difference between the price of the method stated above with the Black-Scholes European option pricing formula, also known as the analytical solution for a European option.

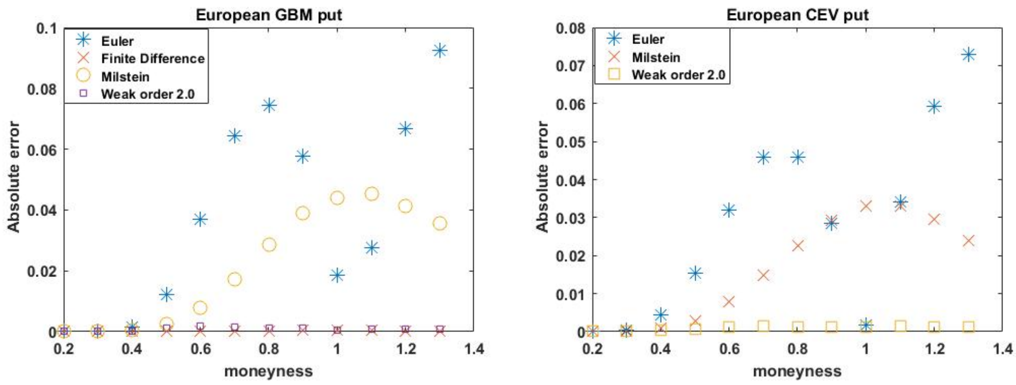

Fig. 3.6: Accuracy of pricing methods for the European put option as a function of moneyness

It is observed that when the option is struck far out-the-money, the prices generated by the four methods are similar in accuracy but as the strike increases, the RMQ Euler-Maruyama scheme starts to exhibit a variable degree of accuracy compared to the other methods. Interestingly in both the CEV and the GBM diffusion processes, the RMQ Euler-Maruyama scheme shows a high degree of accuracy around the at-the-money price. As can be seen, both higher order schemes, the RMQ Milstein scheme and the RMQ weak order 2.0 are accurate over the entire range of strikes, but the RMQ weak order 2.0 is the most accurate scheme of these methods.

3.4.2 Bermudan put option price

The following methods use the Backward Dynamic Programming Principle to price the Bermudan option. These methods are the least squares Monte Carlo method, the RMQ Euler-Maruyama scheme, the RMQ Milstein scheme and the RMQ weak order 2.0 Taylor scheme.

Once all the matrices and vectors from the previous chapter are computed (or defined) using the RMQ algorithm, a high-level algorithm for Bermudan option pricing can then be specified as was reviewed by McWalter *et al.* (2017):

1. Initialize $h_K = H(\Gamma_K, X)$.
2. For $k = K-1, \ldots, 1$
   Set $h_k = \max\big(H(\Gamma_k, X), e^{-r\Delta t} P_{k+1} h_{k+1}\big)$
3. Set $H_0 = e^{-r\Delta t} p_1 h_1$,

where $H_0 = e^{-rT}\mathbb{E}\big[H(S_T, X)\big] \approx e^{-rT} p_K H(\Gamma_K, X)$ is the value of the claim at time $t_0 = 0$ and $H(\Gamma_K, X)$ is the payoff function H applied element-wise to Γ_K and X is the strike. The max function is also applied element-wise with the second argument being the continuation value. Since the transition probability matrix is computed in the previous RMQ algorithm steps, the continuation value is easily computed because of the availability of this transition probability matrix at each time step at no extra cost, making the algorithm efficient.

Similar to Figure 3.6 above, Figure 3.7 exhibits the same trend, in that when the option is struck far out-the-money, the prices generated by the four methods are similar in accuracy but the RMQ weak order 2.0 scheme is the most accurate as was found in the European option case, while the RMQ Euler-Maruyama scheme shows the most variable degree of accuracy. Interestingly the RMQ Euler-Maruyama seems to be very accurate around the at-the-money prices as was observed in the European option case.

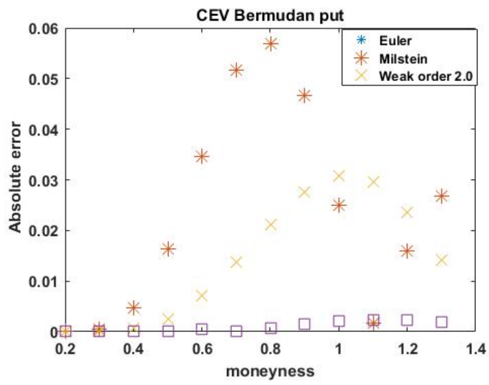

Fig. 3.7: Accuracy of pricing methods for the Bermudan put option as a function of moneyness

Accuracy and computational efficiency

Since the methods are all different, comparing them for efficiency and accuracy means that, a common trait amongst the methods must be compared. To determine accuracy, a high resolution Crank-Nicolson finite difference scheme is used as a proxy for the exact price, the absolute difference between this price and the chosen method, gives the accuracy. To accurately reflect computational efficiency, the accuracy computed must be compared to execution time.

Figure 3.8 shows a plot of the accuracy as a function of execution time, where the cardinality is varied for the RMQ schemes and the sample size is varied for the least squares Monte Carlo method. For the out-the-money case, the variability in accuracy as sample size is increased in the least squares Monte Carlo method occurs as a result of generating independent updates (samples), making this method the least accurate of the four methods. Following the least squares method is the RMQ Euler-Maruyama scheme, this scheme has the shortest execution time as the cardinality is varied but is not the most accurate and efficient. The RMQ weak order 2.0 Taylor scheme is the most accurate and efficient scheme for a given cardinality, but if the cardinality is varied, the execution time increases and the accuracy is also improved. The RMQ Milstein scheme shares similarities with the RMQ weak order 2.0 Taylor scheme in that as the cardinality is increased, execution time is also increased but this scheme is not as accurate as the RMQ weak order 2.0 Taylor scheme. The in-the-money prices show very similar results to the out-the-money prices, the only difference is the accuracy of the RMQ Euler-Maruyama scheme, which is more accurate and more efficient than the RMQ Milstein scheme. The RMQ weak order 2.0 Taylor scheme is the most accurate and efficient method, whilst the least squares Monte Carlo method produces results which are the least and efficient.
Figure 3.9 results are similar to the in-the-money results, in that the RMQ Euler-Maruyama scheme is more accurate and more efficient than the RMQ Milstein scheme. The RMQ weak order 2.0 Taylor scheme is the most accurate and efficient method for a given cardinality but as the cardinality increased, the execution time also increased.

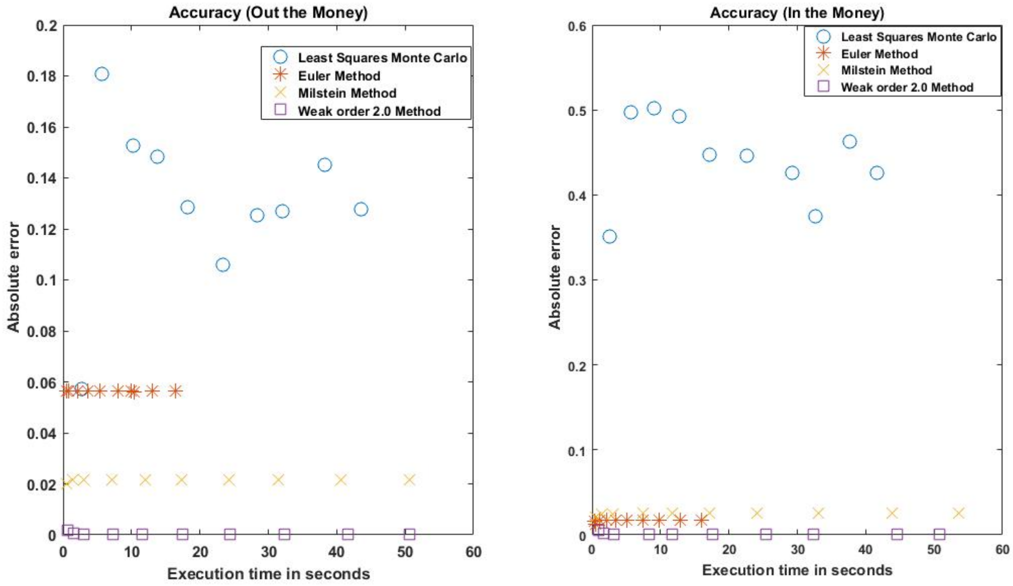

Fig. 3.8: Accuracy as a function of execution times for in-the-money and out-the-money prices

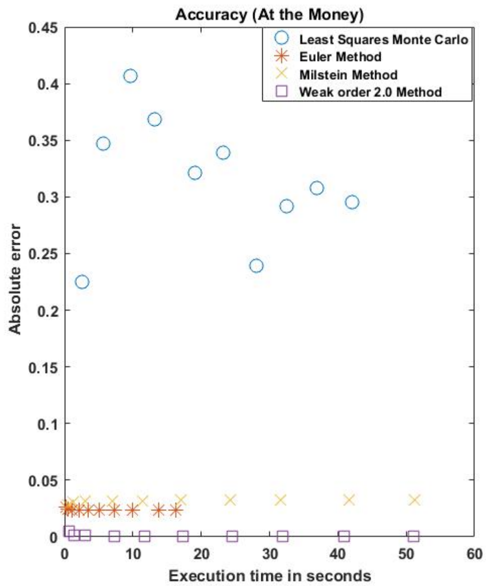

Fig. 3.9: Accuracy as a function of execution times for at-the-money prices

Timing comparison

Lastly, the methods are then compared for timing execution without taking into consideration accuracy. This is to assess the execution time required for each of these methods to price a Bermudan put option. In the table below, all parameters are kept constant and the strike is varied for out-the-money, at-the-money and in-the-money pricing.

Execution time (seconds)			
Method	OTM (K=80)	ATM (K=100)	ITM (K=120)
LSMC	18.6931	18.0845	18.1326
RMQ Euler	3.1387	2.8225	3.4215
RMQ Milstein	9.5864	9.7493	9.5192
RMQ Weak order	9.4673	9.8943	10.0921

Tab. 3.2: Time to execution varying the strike prices.

The results in the table above were taken from a computer with a much slower processing speed, but it can be seen that on a timing comparison, least squares Monte Carlo takes the longest time to compute the Bermudan option price, whilst the RMQ schemes take the shortest time. This can be attributed to how the RMQ algorithms are set up versus the least squares Monte Carlo algorithm. There is also a trade between accuracy and execution time, the RMQ weak order 2.0 scheme takes slightly longer to compute the Bermudan put option price as compared to the RMQ Euler method,but the few seconds that are gained in the RMQ weak order 2.0 scheme are more than compensated for in the accuracy of the method.

Chapter 4

Conclusion

The main focus of the dissertation was to use recursive marginal quantization (RMQ) proposed by Pagès and Sagna (2015) and extended by McWalter *et al.* (2017) for higher order schemes to price a Bermudan option and compare this to other existing methods like the least squares Monte Carlo method and finite difference method to see how computationally efficient and accurate the method is.

A Bermudan put option was priced using three methods: least squares Monte Carlo method, finite difference method and recursive marginal quantization. A high resolution Crank-Nicolson finite difference method was used as a proxy for the exact solution, as Bermudan options do not have a closed-form solution like the European put option. Then the method of Pagès and Sagna (2015) and the extended work by McWalter *et al.* (2017) for higher order schemes was implemented.

In terms of accuracy, the least squares Monte Carlo method was found to be the least accurate method for pricing a Bermudan option because of its variability when generating independent updates as sample size increased. The RMQ Euler-Maruyama scheme was much more accurate than the least squares method, but the higher order RMQ weak order 2.0 scheme proved to be the most accurate. The RMQ Milstein scheme showed a varied result, it was found to be more accurate than the RMQ Euler-Maruyama scheme for the out-the-money pricing but both schemes showed similar accuracy for at-the-money and in-the-money pricing, but this may be as a result of the chosen strikes. Generally, the RMQ higher order schemes showed a better degree of accuracy as a result of the improved approximation of the marginal distributions given by these higher order schemes.

In terms of computational efficiency, which was defined as time to execution versus accuracy, the least squares Monte Carlo method was found to be the least computationally efficient method. The Euler-Maruyama scheme was found to be computationally efficient given loose accuracy constraints. The RMQ weak order 2.0 Taylor scheme was found to be the most computationally efficient method, given the degree of accuracy and time to execution. The RMQ Milstein scheme showed a very similar pattern to the RMQ weak order 2.0 Taylor scheme in terms computational efficiency. Generally, the computational efficiency of the higher order schemes is as a result of the generalized simple matrix formulation which allowed for efficient implementation of these higher order schemes.

In the last section of the previous chapter, a timing comparison was done across different strikes. The least squares Monte Carlo method took the longest time to compute a Bermudan put option price, this is as a result of how the independent updates are generated in the algorithm. The RMQ Euler-Maruyama scheme took the shortest time to generate a Bermudan put option price, but this speed is seen to conflict with the accuracy of the method. The higher order schemes, RMQ weak order 2.0 Taylor scheme and the Milstein scheme, have similar times which are slightly longer than the RMQ Euler-Maruyama scheme, but this trade is then well compensated for in the accuracy of the schemes.

In this dissertation, a Bermudan option under the constant elasticity of variance (CEV) model was priced and it was found that the recursive marginal quantization (RMQ) method was the most computationally efficiency and accurate method.

Bibliography

Barone-Adesi, G. and Whaley, R. E. (1987). Efficient analytic approximation of American option values, *The Journal of Finance* **42**(2): 301–320.

Bekaert, G. and Wu, G. (2000). Asymmetric volatility and risk in equity markets, *Review of Financial Studies* **13**(1): 1–42.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *The Journal of Political Economy* pp. 637–654.

Brennan, M. J. and Schwartz, E. S. (1978). Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis, *Journal of Financial and Quantitative Analysis* **13**(03): 461–474.

Broadie, M. and Detemple, J. (1996). American option valuation: new bounds, approximations, and a comparison of existing methods, *Review of Financial Studies* **9**(4): 1211–1250.

Bru, B. and Yor, M. (2002). Comments on the life and mathematical legacy of wolfgang doeblin, *Finance and Stochastics* **6**(1): 3–47.

Carr, P. (1998). Randomization and the American put, *Review of Financial Studies* **11**(3): 597–626.

Carr, P., Jarrow, R. and Myneni, R. (1992). Alternative characterizations of American put options, *Mathematical Finance* **2**(2): 87–106.

Cox, J. (1975). Notes on option pricing i: constant elasticity of variance diffusions, *Unpublished note, Stanford University, Graduate School of Business* .

Cox, J. C. and Ross, S. A. (1976). The valuation of options for alternative stochastic processes, *Journal of Financial Economics* **3**(1-2): 145–166.

Cox, J. C., Ross, S. A. and Rubinstein, M. (1979). Option pricing: A simplified approach, *Journal of Financial Economics* **7**(3): 229–263.

Delbaen, F. and Shirakawa, H. (2002). A note on option pricing for the constant elasticity of variance model, *Asia-Pacific Financial Markets* **9**(2): 85–99.

Dennis, P. and Mayhew, S. (2002). Risk-neutral skewness: Evidence from stock options, *Journal of Financial and Quantitative Analysis* **37**(03): 471–493.

El Karoui, N. and Karatzas, I. (1995). The optimal stopping problem for a general American put-option, *IMA Volumes in Mathematics and its Applications* **65**: 63–63.

Geske, R. and Johnson, H. E. (1984). The American put option valued analytically, *The Journal of Finance* **39**(5): 1511–1524.

Hull, J. C. (2006). *Options, futures, and other derivatives*, Pearson Education India.

Jacka, S. . (1991). Optimal stopping and the American put, *Mathematical Finance* **1**(2): 1–14.

Jamshidian, F. (1989). Formulas for American options, *Merrill Lynch Capital Markets working paper* .

Johnson, H. (1983). An analytical approximation for the American put price, *Journal of Financial and Quantitative Analysis* **18**(1): 141–148.

Kim, I. J. (1990). The analytic valuation of American options, *Review of Financial Studies* **3**(4): 547–572.

Longstaff, F. A. and Schwartz, E. S. (2001). Valuing American options by simulation: a simple least-squares approach, *Review of Financial Studies* **14**(1): 113–147.

MacMillan, L. (1986). Analytical approximation for the American put option, advances in futures and options research, *Advances in Futures and Options Research* **1**(1): 119–139.

McKean, H. P. (1965). Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics, *Sloan Management Review* **6**(2): 32.

McWalter, T. A. (2016). *Numerical Methods II, class notes*, UCT.

McWalter, T. A., Rudd, R., Kienitz, J. and Platen, E. (2017). Recursive marginal quantization of higher-order schemes, *SSRN net* pp. 1–26.

Pagès, G. and Sagna, A. (2015). Recursive marginal quantization of the Euler scheme of a diffusion process, *Applied Mathematical Finance* **22**(5): 463–498.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula, *the Journal of Finance* **44**(1): 211–219.

Schwartz, E. S. (1977). The valuation of warrants: Implementing a new approach, *Journal of Financial Economics* **4**(1): 79–93.

Strikwerda, J. C. (2004). *Finite difference schemes and partial differential equations*, SIAM.

Sullivan, M. A. (2000). Valuing American put options using Gaussian quadrature, *Review of Financial Studies* **13**(1): 75–94.

Tompaidis, S. and Yang, C. (2014). Pricing American-style options by Monte Carlo simulation: alternatives to ordinary least squares, *SSRN net* .